

On the variational behaviour of functions with positive steepest descent rate

A. Uderzo

October 10, 2014

Abstract This paper investigates some aspects of the variational behaviour of nonsmooth functions, with special emphasis on certain stability phenomena. Relationships linking such properties as sharp minimality, superstability, error bound and sufficiency of first-order optimality conditions are discussed. Their study is performed by employing the steepest descent rate, a rather general tool, which is adequate for a metric space analysis. The positivity of the steepest descent rate is then characterized in terms of Φ -subdifferentials. If specialized to a Banach space setting, the resulting characterizations subsume known results on the stability of error bounds.

Keywords Steepest descent rate · Strong slope · Sharp minimizer · Nondifferentiable optimization · Superstable solution · Error bound · Optimality condition · Φ -subdifferential

Mathematics Subject Classification (2010) 49J52 · 49J53 · 90C25 · 90C31

1 Introduction

It is well known that classical differential calculus provides a powerful apparatus for a refined analysis of optimization problems. On the other hand, successful approaches to the differentiability of functions have revealed intriguing connections between minimality and smoothness, with the result that the latter can be established through variational principles. The author of the present paper shares the opinion of all those believing that the theoretical

A. Uderzo
Dept. of Mathematics and Applications
University of Milano-Bicocca
Tel.: +39-02-64485871
Fax: +39-02-64485705
E-mail: amos.uderzo@unimib.it

framework emerging from this proficuous interplay, active since several centuries, should not exclude nonsmoothness. In fact, historically, the absence of differentiability, when observed, was very often perceived as a pathology (a “miserable plague”, in the Hermite’s words [1]) to be accurately avoided, whenever possible. Since theoretical as well as applicative needs show that this is not always possible (or reasonable), an area called nonsmooth analysis was developed with the specific task to treat such a pathology, especially for those problems arising in optimization. The present paper is an attempt to show that nonsmoothness, along with evident drawbacks and limitations, can also afford some benefits in the analysis of optimization problems. This is done by considering the favourable effects of the variational behaviour of functions having a steepest descent rate, which is positive at some point. The positivity of the steepest descent rate is not consistent with the classical differentiability. In spite of this, it is the source of several robustness phenomena having to do with perturbed optimization: namely they relate to the local sharp minimality, the superstability, the error bound property and its stability. A proper general setting where to study the nature of all these phenomena is that of metric spaces, an environment in which it is not clear how to speak of smoothness and differentiability. Nonetheless, nonsmooth analysis has succeeded in devising generalized differential tools, which reveal to be adequate for a metric space analysis.

The contents of this paper are organized as follows. The next Section 2 starts with introducing the basic notion of steepest descent rate, which will be the basic tool of analysis. By means of that the positivity condition (\mathcal{C}) is formulated. Such a condition could be regarded as a manifestation of nonsmoothness in metric spaces. In the subsections included in Section 2 condition (\mathcal{C}) is shown to be equivalent to local sharp minimality, to superstability of a solution to a perturbed optimization problem and, to a certain extent, to the local error bound. What is more, in the presence of (\mathcal{C}) the last property turns out to be stable with respect to perturbations with controlled strong slope. Section 3 is devoted to the characterization of the positivity condition (\mathcal{C}) in terms of global and local Φ -subdifferentials. In Section 4 the findings of the previous section are specialized to a Banach space setting, where widely employed nonsmooth analysis tools such as the Hadamard generalized derivative, the regular subdifferential, the subdifferential in the sense of convex analysis can be utilized. Relationships with existing results from the related literature are discussed. A final section is reserved to distil the spirit of the analysis here exposed.

2 Steepest descent rate and variational analysis in metric spaces

Whenever $r \in \mathbb{R} \cup \{\pm\infty\}$, symbol $[r]_+$ stands for $\max\{r, 0\}$. Given a metric space (X, d) , the closed ball with center $x \in X$ and radius $r \geq 0$ is denoted by $B(x, r)$. If $x \in X$ and $S \subseteq X$, the distance of x from S is indicated by $\text{dist}(x, S) = \inf_{y \in S} d(x, y)$, with the convention that $\text{dist}(x, \emptyset) = +\infty$. Given

a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$, its domain is denoted by $\text{dom } f = \{x \in X : |f(x)| < \infty\}$. If $\alpha \in \mathbb{R}$, $[f \leq \alpha] = \{x \in X : f(x) \leq \alpha\}$ and $[f > \alpha] = X \setminus [f \leq \alpha]$ indicate the α -sublevel set and the strict α -superlevel set of f , respectively. In particular, whenever it is $\inf_X f > -\infty$, $\text{Argmin}(f) = [f \leq \inf_X f]$ denotes the set of all global minimizers of f , if any. Throughout the paper, the acronym l.s.c. stands for lower semicontinuous.

The analysis of the variational properties of functions in metric spaces will be mainly conducted by means of the following basic tool.

Definition 1 Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ defined on a metric space X and an element $\bar{x} \in \text{dom } f$, the value

$$f^\downarrow(\bar{x}) = \liminf_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})},$$

is called the *steepest descent rate* of f at \bar{x} .

To the best of the author's knowledge, the first employment of the steepest descent rate in connection with extremum problems in metric spaces goes back to [2], where the notion of inf-stationary point is introduced. Later on it found relevant applications in nonsmooth analysis (see, for instance, [3]). By his side, V.F. Demyanov contributed to popularize the use of this tool as well as of its k -th order version: in several of his works he employed it for formulating optimality conditions in metric spaces, as a starting point for further developments in nondifferentiable optimization (see, for instance, [4, 5, 6, 7], wherefrom the notation of Definition 1 has been borrowed). Further recent employments can be found in [8, 9]. Clearly the steepest descent rate is strictly connected with another tool, widely utilized in metric space variational analysis, known as strong slope. Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and an element $\bar{x} \in \text{dom } f$, according to [10], by the strong slope (or calmness rate) of f at \bar{x} the value

$$|\nabla f|(\bar{x}) = \begin{cases} 0, & \text{if } \bar{x} \text{ is a local minimizer for } f, \\ \limsup_{x \rightarrow \bar{x}} \frac{f(\bar{x}) - f(x)}{d(x, \bar{x})}, & \text{otherwise} \end{cases}$$

is meant. Now, it is readily seen that if $\bar{x} \in \text{dom } f$ is a local minimizer of f , then $f^\downarrow(\bar{x}) \geq 0$. Such a condition is evidently only necessary for the local minimality of \bar{x} . Nevertheless, its enforcement

$$(C) \quad f^\downarrow(\bar{x}) > 0$$

is a sufficient condition for local (strict) optimality (see [4]). One of the aims of this article is to show that actually condition (C) can tell even more than that. Notice that, whenever (C) holds, it has to be $|\nabla f|(\bar{x}) = 0$. In circumstances in which the annihilating of the strong slope does not allow to guarantee those benefits deriving from nondegeneracy conditions, condition (C) turns out to provide meaningful insights into the variational behaviour of f near \bar{x} . As it will be illustrated in Sect. 4, when functions are defined in more structured spaces, the occurrence of (C) is essentially connected with the nonsmoothness of f .

2.1 Sharp minimizers and their superstability

The next definition, originally introduced in [11] in its global form within the context of nondifferentiable convex optimization, captures a possible variational behaviour of a function near a local minimizer of it. It describes how the local minimum value is attained at that point.

Definition 2 Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$, an element $\bar{x} \in \text{dom } f$ is said to be a *local sharp minimizer* of f if there exist positive σ and r such that

$$f(x) \geq f(\bar{x}) + \sigma d(x, \bar{x}), \quad \forall x \in B(\bar{x}, r). \quad (1)$$

The value

$$\text{sha}(f, \bar{x}) = \sup\{\sigma > 0 : \exists r > 0 \text{ satisfying (1)}\}$$

will be called *modulus of local sharpness* of f at \bar{x} . If inequality (1) continues being true with $B(\bar{x}, r)$ replaced by X , then \bar{x} is called *global sharp minimizer* of f .

Example 1 Here some simple situations are illustrated in which the notion of sharp minimality naturally arises.

(i) Let (X, d) be a complete metric space and let $T : X \rightarrow X$ be a contraction mapping, i.e. there exists $\alpha \in [0, 1)$ such that $d(T(x_1), T(x_2)) \leq \alpha d(x_1, x_2)$, for every $x_1, x_2 \in X$. The Banach-Caccioppoli fixed point theorem ensures the existence of a unique fixed point $\bar{x} \in X$, such that

$$d(x, \bar{x}) \leq \frac{1}{1-\alpha} d(x, T(x)), \quad \forall x \in X.$$

This inequality shows that the displacement function $f_T : X \rightarrow [0, +\infty)$ of T , defined as $f_T(x) = d(x, T(x))$, admits \bar{x} as a global sharp minimizer, with $\sigma = 1 - \alpha$.

(ii) Whenever $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ admits $\bar{x} \in X$ as a local sharp minimizer and $\tilde{f} : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ is such that

$$\tilde{f}(\bar{x}) = f(\bar{x}) \quad \text{and} \quad \tilde{f}(x) \geq f(x), \quad \forall x \in B(\bar{x}, r),$$

for some $r > 0$, \bar{x} is a local sharp minimizer also for \tilde{f} . So, if $X = \mathbb{R}^n$ and $\|\cdot\|$ stands for the Euclidean norm, suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is any radial function with profile $\pi_f : [0, +\infty) \rightarrow [0, +\infty)$ satisfying the inequality

$$\pi_f(t) \geq \sigma t, \quad \forall t \in [0, r]$$

for some $r, \sigma > 0$, and let $\bar{x} \in \mathbb{R}^n$. Then, every function $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$\tilde{f}(x) \geq f(x - \bar{x}) + c, \quad \forall x \in B(\bar{x}, r),$$

with $c \in \mathbb{R}$, admits \bar{x} as a local sharp minimizer. Moreover, it results in $\text{sha}(f, \bar{x}) \geq \sigma$.

Notice that a local sharp minimizer is a strict minimizer of f and an isolated point of the sublevel set $[f \leq f(\bar{x})]$. As an immediate consequence of Definition 2, one obtains the following local form of Tykhonov well-posedness: if $(x_n)_{n \in \mathbb{N}}$ is a sequence in X , whose elements lie sufficiently near \bar{x} , and $f(x_n) \rightarrow f(\bar{x})$ as $n \rightarrow \infty$, then it must be $x_n \rightarrow \bar{x}$, as $n \rightarrow \infty$. Of course, if \bar{x} is a global sharp minimizer of f , the extremum problem $\min_{x \in X} f$ is Tykhonov well-posed.

It was remarked already in [11] that condition (1) can not be satisfied a priori by smooth functions, if considered in a properly structured setting. In spite of this, sharp minimality ensures good properties. For instance, it has been shown that sharp minimality is a sufficient condition for finite termination of the proximal point algorithm (see [12, 13]). A weaker version of the notion sharp minimality (known as weak sharp minimality) gained an even major success, due to its recognized relevance in the convergence analysis of algorithms for solving extremum problems as well as in the study of stability of variational problems (see [14, 15]).

Local sharpness of minimizers can be easily characterized in terms of positivity of the steepest descent rate, as follows.

Proposition 1 *Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$, an element $\bar{x} \in \text{dom } f$ is a local sharp minimizer iff condition (C) holds. Moreover $f^\downarrow(\bar{x}) = \text{sha}(f, \bar{x})$.*

Proof The proof of the first assertion is a straightforward consequence of Definition 1 and Definition 2. To see the inequality $f^\downarrow(\bar{x}) \geq \text{sha}(\bar{x})$, fix an arbitrary $\sigma > 0$ such that (1) holds for some $r > 0$. One has

$$f^\downarrow(\bar{x}) = \sup_{\delta > 0} \inf_{x \in B(\bar{x}, \delta) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} \geq \inf_{x \in B(\bar{x}, r) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} \geq \sigma.$$

On the other hand, by Definition 1, for any arbitrary $\epsilon > 0$ there exists $r_\epsilon > 0$ such that

$$\inf_{x \in B(\bar{x}, r_\epsilon) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} > f^\downarrow(\bar{x}) - \epsilon,$$

so inequality (1) is satisfied by $\sigma = f^\downarrow(\bar{x}) - \epsilon$ for $r = r_\epsilon$. If it were $f^\downarrow(\bar{x}) > \text{sha}(f, \bar{x})$, by taking $\epsilon < f^\downarrow(\bar{x}) - \text{sha}(f, \bar{x})$ this would contradict the definition of $\text{sha}(f, \bar{x})$. \square

The variational behaviour characterized by condition (C) yields favorable stability properties of solutions in perturbed optimization. Let us consider indeed the following family of perturbed problems:

$$(\mathcal{P}_g) \quad \min_{x \in X} [f(x) + g(x)]$$

where $g \in \mathcal{G}_{\bar{x}}$ and

$$\mathcal{G}_{\bar{x}} = \{g : X \rightarrow \mathbb{R} \cup \{\pm\infty\} : \bar{x} \in \text{dom } g\}.$$

Notice that the additive perturbation term g allows one to cover very general perturbation effects. Clearly, when $g \equiv 0$ one gets the unperturbed problem $\min_{x \in X} f(x)$. The next definition generalizes a strong concept of stability in optimization, which was proposed again in [11].

Definition 3 With reference to a family of perturbed problems (\mathcal{P}_g) , a local solution $\bar{x} \in \text{dom } f$ of (\mathcal{P}_0) is called *superstable* (for (\mathcal{P}_g)) if there exists $\epsilon_0 > 0$ such that \bar{x} locally solves (\mathcal{P}_g) , for every $g \in \mathcal{G}_{\bar{x}}$, with $|\nabla g|(\bar{x}) < \epsilon_0$.

Remark 1 As a comment to Definition 3, observe that condition $|\nabla g|(\bar{x}) < \epsilon_0$ holds in particular, whenever $g \in \mathcal{G}_{\bar{x}}$ is locally Lipschitz around \bar{x} , with Lipschitz constant

$$\text{lip}(g, \bar{x}) = \limsup_{\substack{x_1, x_2 \rightarrow \bar{x} \\ x_1 \neq x_2}} \frac{|g(x_1) - g(x_2)|}{d(x_1, x_2)} < \epsilon_0.$$

Thus, if \bar{x} is superstable, it persists as a solution to (\mathcal{P}_g) under locally Lipschitz perturbations of f . More generally, if \bar{x} is superstable it locally solves any problem $\min_{x \in X} \tilde{f}(x)$, for every $\tilde{f} \in \text{Ptb}(f, \bar{x}, \epsilon_0)$, where

$$\text{Ptb}(f, \bar{x}, \epsilon_0) = \left\{ \tilde{f} \in \mathcal{G}_{\bar{x}} : \limsup_{x \rightarrow \bar{x}} \frac{|\tilde{f}(x) - f(x) - (\tilde{f}(\bar{x}) - f(\bar{x}))|}{d(x, \bar{x})} \leq \epsilon_0 \right\}.$$

Indeed, if $\tilde{f} \in \text{Ptb}(f, \bar{x}, \epsilon_0)$, then

$$|\nabla(\tilde{f} - f)|(\bar{x}) \leq \limsup_{x \rightarrow \bar{x}} \frac{|\tilde{f}(x) - f(x) - (\tilde{f}(\bar{x}) - f(\bar{x}))|}{d(x, \bar{x})}.$$

The above kind of perturbations has been already considered in [16] and will be again employed here in a subsequent section.

The next proposition reveals that the superstability property, as presented in Definition 3, is actually a reformulation of the local sharp minimality.

Proposition 2 Let $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ be a given function and let $\bar{x} \in \text{dom } f$. Then, \bar{x} is a local sharp minimizer iff it is superstable for (\mathcal{P}_g) .

Proof Let us start with supposing \bar{x} to be a local sharp minimizer of f . Then condition (\mathcal{C}) does hold. So, take $\epsilon_0 = f^\downarrow(\bar{x})$. For any $g \in \mathcal{G}_{\bar{x}}$, with $|\nabla g|(\bar{x}) < \epsilon_0$ one obtains

$$(f + g)^\downarrow(\bar{x}) \geq \liminf_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} + \liminf_{x \rightarrow \bar{x}} \frac{g(x) - g(\bar{x})}{d(x, \bar{x})} \geq f^\downarrow(\bar{x}) - |\nabla g|(\bar{x}) > 0.$$

Since (\mathcal{C}) is a sufficient optimality condition, this implies that \bar{x} is also a local minimizer of $f + g$, for every $g \in \mathcal{G}_{\bar{x}}$, with $|\nabla g|(\bar{x}) < \epsilon_0$.

Suppose now that \bar{x} satisfies Definition 3 with some $\epsilon_0 > 0$. Choose $\epsilon \in (0, \epsilon_0)$ and observe that the function $g : X \rightarrow \mathbb{R}$ defined by $g(x) = -\epsilon d(x, \bar{x})$ belongs to $\mathcal{G}_{\bar{x}}$. Moreover, one sees that in such case it is $|\nabla g|(\bar{x}) = \epsilon$. Thus for some $r > 0$ it must hold

$$f(x) + g(x) = f(x) - \epsilon d(x, \bar{x}) \geq f(\bar{x}) + g(\bar{x}), \quad \forall x \in B(\bar{x}, r),$$

which allows one to conclude that \bar{x} is a local sharp minimizer of f . This completes the proof. \square

Remark 2 It is worth noting that the proof of Proposition 2 actually reveals that if condition (C) is valid for f at \bar{x} , it continues being valid for any perturbed function $f+g$ at the same point, for every $g \in \mathcal{G}_{\bar{x}}$, with $|\nabla g|(\bar{x}) < f^\downarrow(\bar{x})$. In other words, sharp minimality itself is stable under this kind of perturbation.

2.2 Sufficiency in optimality conditions

We have seen that the steepest descent rate enables one to express a sufficient condition for local optimality. The next proposition shows how the same notion can be employed in formulating a sufficient condition for the (global) solution existence.

Proposition 3 *Let (X, d) be a complete metric space and let $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function with $\inf_X f > -\infty$. If there exists $\sigma > 0$ such that*

$$\sup_{x \in [f > \inf_X f] \cap \text{dom } f} f^\downarrow(x) < -\sigma, \quad (2)$$

then $\text{Argmin}(f) \neq \emptyset$.

Proof If $f \equiv +\infty$ the thesis is trivially true as it is $\text{Argmin}(f) = X$. Otherwise, take an element $x_0 \in X$ such that $f(x_0) < \inf_X f + \sigma$. Since f is l.s.c. and bounded from below, and X is metrically complete by hypothesis, it is possible to invoke the Ekeland variational principle. According to it, there exists $\bar{x} \in B(x_0, 1)$ such that $f(\bar{x}) \leq f(x_0)$, so $\bar{x} \in \text{dom } f$, and

$$f(\bar{x}) < f(x) + \sigma d(x, \bar{x}), \quad \forall x \in X \setminus \{\bar{x}\}. \quad (3)$$

Suppose now that $\bar{x} \in [f > \inf_X f]$. Then one finds as a consequence of inequality (3) that $f^\downarrow(\bar{x}) \geq -\sigma$, which contradicts the hypothesis (2). Therefore it must be $\bar{x} \in [f \leq \inf_X f]$, so \bar{x} turns out to be a global minimizer of f . \square

2.3 Error bound for inequalities

The notion of local/global error bound is known to play a key role in optimization and variational analysis. Among other topics, it emerges as a crucial concept in deriving exact penalty functions of constrained optimization problems (see [17], Ch. 6.8) as well as in connection with the property of calmness (equivalently, metric subregularity) (see, for instance, [18, 19, 20]).

Definition 4 Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and an element $\bar{x} \in X$, with $f(\bar{x}) = 0$, f is said to admit a *local error bound* at \bar{x} if there exist reals $c > 0$ and $r > 0$ such that

$$\text{dist}(x, [f \leq 0]) \leq c[f(x)]_+, \quad \forall x \in B(\bar{x}, r). \quad (4)$$

If inequality (4) continues to hold with $B(\bar{x}, r)$ replaced by X , f is said to admit a *global error bound* at \bar{x}

Remark 3 As done for instance in [16], it is worth observing that the best (lower) bound of all constants c for which inequality (4) is true coincides with the value $(\text{Er}f(\bar{x}))^{-1}$, where

$$\text{Er}f(\bar{x}) = \liminf_{\substack{x \rightarrow \bar{x} \\ f(x) > 0}} \frac{f(x)}{\text{dist}(x, [f \leq 0])}$$

is called the error bound modulus (aka conditioning rate) of f at \bar{x} .

Proposition 4 *Let $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$ be such that $f(\bar{x}) = 0$. If condition (C) holds, then f admits a local error bound at \bar{x} . Moreover, it holds*

$$f^\downarrow(\bar{x}) \leq \text{Er}f(\bar{x}).$$

Proof By virtue of condition (C), \bar{x} is a strict local minimizer of f . More precisely, for every $\epsilon \in (0, f^\downarrow(\bar{x}))$, there exists $\delta_\epsilon > 0$ such that

$$f(x) \geq (f^\downarrow(\bar{x}) - \epsilon)d(x, \bar{x}) > 0, \quad \forall x \in B(\bar{x}, \delta_\epsilon) \setminus \{\bar{x}\}. \quad (5)$$

This entails that

$$[f \leq 0] \cap B(\bar{x}, \delta_\epsilon) = \{\bar{x}\}.$$

Consequently, one finds

$$\text{dist}(x, [f \leq 0]) = d(x, \bar{x}), \quad \forall x \in B(\bar{x}, \delta_\epsilon/2),$$

whence, taking account of inequality (5), it readily follows

$$\text{dist}(x, [f \leq 0]) \leq (f^\downarrow(\bar{x}) - \epsilon)^{-1}[f(x)]_+, \quad \forall x \in B(\bar{x}, \delta_\epsilon/2).$$

This proves that f admits a local error bound at \bar{x} . Besides, since $(\text{Er}f(\bar{x}))^{-1}$ is the lower bound of all constants c satisfying inequality (4), it follows

$$f^\downarrow(\bar{x}) - \epsilon \leq \text{Er}f(\bar{x}).$$

The arbitrariness of ϵ allows one to conclude the proof. \square

Through simple counterexamples, one quickly realizes that the local error bound property can take place even if $f^\downarrow(\bar{x}) = 0$, namely condition (C) is in general only sufficient for it. Nonetheless, if \bar{x} strictly minimizes f , condition (C) becomes also necessary.

Proposition 5 *Let $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$ be such that $f(\bar{x}) = 0$. If \bar{x} is a local strict minimizer of f and f admits a local error bound at \bar{x} , then condition (C) holds true.*

Proof By the local error bound assumption at \bar{x} , one has that for some $\delta, c > 0$ it is

$$\text{dist}(x, [f \leq 0]) \leq c[f(x)]_+, \quad \forall x \in B(\bar{x}, \delta).$$

Since $f(\bar{x}) = 0$ and \bar{x} is a strict local minimizer of f , by a proper reduction of the value of δ , one has

$$f(x) > 0, \quad \forall x \in B(\bar{x}, \delta).$$

Consequently, it results in

$$[f(x)]_+ = f(x) \quad \text{and} \quad \text{dist}(x, [f \leq 0]) = d(x, \bar{x}), \quad \forall x \in B(\bar{x}, \delta/2).$$

It follows

$$\inf_{x \in B(\bar{x}, \delta/2) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} \geq \frac{1}{c},$$

whence one obtains the thesis. \square

While in general condition (C) can not characterize the validity of the local error bound property for f at \bar{x} , it enables one to single out a stronger property than that: actually, it guarantees the local error bound property to hold for the whole family of functions, which are sufficiently small perturbations of f , in a sense clarified by the next proposition.

Corollary 1 *Let $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$ be such that $f(\bar{x}) = 0$. If condition (C) holds, then $f + g$ admits a local error bound at \bar{x} , for every $g \in \mathcal{G}_{\bar{x}}$ such that $g(\bar{x}) = 0$ and $|\nabla g|(\bar{x}) < f^\downarrow(\bar{x})$, and it results in*

$$f^\downarrow(\bar{x}) - |\nabla g|(\bar{x}) \leq \text{Er}(f + g)(\bar{x}).$$

Proof As noticed in Remark 2, condition (C) is stable under additive perturbation of f by $g \in \mathcal{G}_{\bar{x}}$, provided that $|\nabla g|(\bar{x}) < f^\downarrow(\bar{x})$. Then, it suffices to apply Proposition 4 to function $f + g$ at \bar{x} and to recall that $(f + g)^\downarrow(\bar{x}) \geq f^\downarrow(\bar{x}) - |\nabla g|(\bar{x})$. \square

Remark 4 It is worth noting that, under the hypotheses of Corollary 1, the local error bound at \bar{x} remains in force for every $\tilde{f} \in \text{Ptb}(f, \bar{x}, \epsilon)$, with $\epsilon < f^\downarrow(\bar{x})$.

3 Subdifferential characterizations of condition (C)

In what follows, given a metric space (X, d) , $\text{Lip}(X)$ will denote the vector space of all Lipschitz continuous functionals defined on X , equipped with the quasinorm

$$\|\phi\|_{\text{Lip}} = \sup_{\substack{x_1, x_2 \in X \\ x_1 \neq x_2}} \frac{|\phi(x_1) - \phi(x_2)|}{d(x_1, x_2)}, \quad \phi \in \text{Lip}(X),$$

which induces the quasimetric $d_{\text{Lip}} : \text{Lip}(X) \times \text{Lip}(X) \rightarrow [0, +\infty)$. If introducing the equivalence relation over $\text{Lip}(X)$

$$\phi_1 \sim \phi_2 \quad \text{iff} \quad \exists c \in \mathbb{R} : \phi_1(x) - \phi_2(x) = c, \quad \forall x \in X,$$

then $\|\cdot\|_{\text{Lip}}$ is well defined on equivalence classes and $(\text{Lip}(X)/\sim, \|\cdot\|_{\text{Lip}})$ turns out to be a Banach space (see [21]). Once fixed a nonempty family $\Phi \subseteq \text{Lip}(X)/\sim$, it is possible to introduce related concepts of subgradients and subdifferentials of functions defined on X , as done for instance in [21] (see also references therein). For the sake of notational simplicity, elements in $\text{Lip}(X)/\sim$ will be henceforth indicated with the same symbols as their representatives in $\text{Lip}(X)$.

Definition 5 Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$, $\bar{x} \in \text{dom } f$, and $\Phi \subseteq \text{Lip}(X)/\sim$, the set

$$\partial_{\Phi} f(\bar{x}) = \{\phi \in \Phi : f(x) - f(\bar{x}) \geq \phi(x) - \phi(\bar{x}), \quad \forall x \in X\}$$

is called the Φ -subdifferential of f at \bar{x} ; again, the set

$$\partial_{\Phi}^{\text{loc}} f(\bar{x}) = \bigcap_{\epsilon > 0} \partial_{\Phi, \epsilon}^{\text{loc}} f(\bar{x}),$$

where, for a given $\epsilon \geq 0$, it is

$$\partial_{\Phi, \epsilon}^{\text{loc}} f(\bar{x}) = \{\phi \in \Phi : \exists r > 0 : f(x) - f(\bar{x}) \geq \phi(x) - \phi(\bar{x}) - \epsilon d(x, \bar{x}), \\ \forall x \in B(\bar{x}, r)\},$$

is called the *local* Φ -subdifferential of f at \bar{x} .

Remark 5 Take into account that the inequalities defining the Φ -subdifferential and its local counterpart involve an abuse of notation, because ϕ should indicate a class and not a single function. Nevertheless, since $\phi_1 \sim \phi_2$ implies $\phi_1(x) - \phi_1(\bar{x}) = \phi_2(x) - \phi_2(\bar{x})$ for every $x \in X$, such inequalities are well defined.

From Definition 5 it is readily seen that

$$\partial_{\Phi} f(\bar{x}) \subseteq \partial_{\Phi, 0}^{\text{loc}} f(\bar{x}) \subseteq \partial_{\Phi}^{\text{loc}} f(\bar{x}) \subseteq \partial_{\Phi, \epsilon}^{\text{loc}} f(\bar{x}), \quad \forall \epsilon > 0.$$

Notice that, if equipped with the metric d_{Lip} , Φ becomes a metric space. In the next proposition the topological notion of interior refers to such a metric structure on Φ . The null element of $\text{Lip}(X)$ and its \sim -equivalent class is denoted here by $\mathbf{0}$.

Proposition 6 Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, let $\Phi \subseteq \text{Lip}(X)/\sim$ be such that $\mathbf{0} \in \Phi$. If condition (C) holds then

$$\mathbf{0} \in \text{int } \partial_{\Phi, 0}^{\text{loc}} f(\bar{x}).$$

In particular, for every $\sigma \in (0, f^{\downarrow}(\bar{x}))$, one has $B(\mathbf{0}, \sigma) \subseteq \partial_{\Phi, 0}^{\text{loc}} f(\bar{x})$.

Proof According to condition (C), fixing an arbitrary $\sigma \in (0, f^\downarrow(\bar{x}))$, there exists $r > 0$ such that

$$\frac{f(x) - f(\bar{x})}{d(x, \bar{x})} \geq \sigma, \quad \forall x \in B(\bar{x}, r). \quad (6)$$

So, take an arbitrary $\phi \in B(\mathbf{0}, \sigma)$. Recalling the definition of d_{Lip} , one has in particular

$$\frac{|\phi(x) - \phi(\bar{x})|}{d(x, \bar{x})} \leq \sigma, \quad \forall x \in X \setminus \{\bar{x}\},$$

whence, owing to inequality (6), it is

$$f(x) - f(\bar{x}) \geq \phi(x) - \phi(\bar{x}), \quad \forall x \in B(\bar{x}, r).$$

The last inequality shows that $\phi \in \partial_{\Phi, 0}^{\text{loc}} f(\bar{x})$ and such an inclusion implies that $B(\mathbf{0}, \sigma) \subseteq \partial_{\Phi, 0}^{\text{loc}} f(\bar{x})$. \square

The reader should observe that the above necessary condition holds upon a rather general assumption on the class Φ . It is not difficult to realize, through proper choices of Φ , that without additional assumptions the assertion of Proposition 6 can not be reversed. The next definition is aimed at introducing some additional assumptions on Φ that seem to work in order to formulate sufficient conditions for (C).

Definition 6 Let $\Phi \subseteq \text{Lip}(X)/\sim$ be a family containing $\mathbf{0}$. Φ is said to satisfy the

(i) *supporting distance property* if there exists $\kappa \in (0, +\infty)$ such that for every $\epsilon > 0$

$$\sup_{\phi \in B(\mathbf{0}, \kappa\epsilon)} [\phi(x) - \phi(\bar{x})] \geq \epsilon d(x, \bar{x}), \quad \forall x, \bar{x} \in X; \quad (7)$$

(ii) *local supporting distance property* at \bar{x} if there exists $\kappa \in (0, +\infty)$ such that for every $\epsilon > 0$ there is $r > 0$ such that

$$\sup_{\phi \in B(\mathbf{0}, \kappa\epsilon)} [\phi(x) - \phi(\bar{x})] \geq \epsilon d(x, \bar{x}), \quad \forall x \in B(\bar{x}, r). \quad (8)$$

Clearly, property (i) implies property (ii) in Definition 6. Moreover, if Φ is a class fulfilling (i) or (ii) and $\tilde{\Phi}$ is any other class such that $\tilde{\Phi} \supseteq \Phi$, then also $\tilde{\Phi}$ does. Some examples of families in $\text{Lip}(X)/\sim$ fulfilling the above properties will be provided in the next section (see Remark 8).

Proposition 7 *Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, let $\Phi \subseteq \text{Lip}(X)/\sim$. If Φ satisfies the supporting distance property, then $\mathbf{0} \in \text{int } \partial_\Phi f(\bar{x})$ implies condition (C).*

Proof By hypothesis there exists $\epsilon > 0$ such that $B(\mathbf{0}, \epsilon) \subseteq \partial_\Phi f(\bar{x})$. By virtue of the supporting distance property, there exists $\kappa \in (0, +\infty)$ such that inequality (7) holds true. Consequently, one finds

$$\begin{aligned} f(x) - f(\bar{x}) &\geq \sup_{\phi \in \partial_\Phi f(\bar{x})} [\phi(x) - \phi(\bar{x})] \geq \sup_{\phi \in B(\mathbf{0}, \epsilon)} [\phi(x) - \phi(\bar{x})] \\ &\geq \frac{\epsilon}{\kappa} d(x, \bar{x}), \quad \forall x \in X, \end{aligned}$$

whence

$$\frac{f(x) - f(\bar{x})}{d(x, \bar{x})} \geq \frac{\epsilon}{\kappa}, \quad \forall x \in X \setminus \{\bar{x}\}. \quad (9)$$

From the last inequality one immediately gets the validity of condition (C). \square

Localizing the notion of subdifferential as well as the supporting distance property allows one to obtain a milder sufficient condition. Nevertheless, the price to be paid for such a refinement of the preceding result is an extra compactness assumption to be taken, which limits the range of application.

Proposition 8 *Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, let $\Phi \subseteq \text{Lip}(X)/\sim$. If Φ satisfies the local supporting distance property at \bar{x} and balls in Φ are compact, then $\mathbf{0} \in \text{int } \partial_\Phi^{\text{loc}} f(\bar{x})$ implies condition (C).*

Proof Assume that $B(\mathbf{0}, \sigma) \subseteq \partial_\Phi^{\text{loc}} f(\bar{x})$, for some $\sigma > 0$. By using the supporting distance property at \bar{x} , one gets the existence of $\kappa > 0$ as in (ii) of Definition 6. Notice that one can assume that $\kappa \geq 1$. Take an arbitrary $\phi_0 \in B(\mathbf{0}, \sigma)$. Since ϕ_0 belongs in particular to $\partial_{\Phi, \frac{\sigma}{4\kappa}}^{\text{loc}} f(\bar{x})$, then there exists $r_{\sigma, \phi_0} > 0$ such that

$$f(x) - f(\bar{x}) \geq \phi_0(x) - \phi_0(\bar{x}) - \frac{\sigma}{4\kappa} d(x, \bar{x}), \quad \forall x \in B(\bar{x}, r_{\sigma, \phi_0}),$$

so that

$$\inf_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi_0(x) - \phi_0(\bar{x})]}{d(x, \bar{x})} \geq -\frac{\sigma}{4\kappa}.$$

Without loss of generality it is possible to assume that $r_{\sigma, \phi_0} < \min\{r, \sigma/4\kappa\}$, where r is as in inequality (8), corresponding to $\epsilon = \sigma/\kappa$. Now, observe that

for every $\phi \in B(\phi_0, r_{\sigma, \phi_0})$ it results in

$$\begin{aligned}
& \inf_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi(x) - \phi(\bar{x})]}{d(x, \bar{x})} \geq \\
& \inf_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi_0(x) - \phi_0(\bar{x})]}{d(x, \bar{x})} - \\
& \sup_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{\phi(x) - \phi(\bar{x}) - [\phi_0(x) - \phi_0(\bar{x})]}{d(x, \bar{x})} \geq \\
& \inf_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi_0(x) - \phi_0(\bar{x})]}{d(x, \bar{x})} - \\
& \sup_{\substack{x_1, x_2 \in X \\ x_1 \neq x_2}} \frac{\phi(x_1) - \phi(x_2) - [\phi_0(x_1) - \phi_0(x_2)]}{d(x_1, x_2)} \geq \\
& -\frac{\sigma}{4\kappa} - d_{\text{Lip}}(\phi, \phi_0) \geq -\frac{\sigma}{2\kappa},
\end{aligned}$$

wherefrom it follows

$$\inf_{\phi \in B(\phi_0, r_{\sigma, \phi_0})} \inf_{x \in B(\bar{x}, r_{\sigma, \phi_0}) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi(x) - \phi(\bar{x})]}{d(x, \bar{x})} \geq -\frac{\sigma}{2\kappa}.$$

The family $\{\text{int } B(\phi_0, r_{\sigma, \phi_0}) : \phi_0 \in B(\mathbf{0}, \sigma)\}$ forms an open covering of $B(\mathbf{0}, \sigma)$, which is a compact set by hypothesis. Therefore, there exist $N \in \mathbb{N}$ and $\phi_1, \dots, \phi_N \in B(\mathbf{0}, \sigma)$ such that the subfamily $\{\text{int } B(\phi_i, r_{\sigma, \phi_i}) : i = 1, \dots, N\}$ is still a covering for $B(\mathbf{0}, \sigma)$. Define

$$r_0 = \min_{i=1, \dots, N} r_{\sigma, \phi_i}.$$

Since for every $\phi \in B(\mathbf{0}, \sigma)$ an index $i_* \in \{1, \dots, N\}$ can be found such that $\phi \in B(\phi_{i_*}, r_{\sigma, \phi_{i_*}})$, then by virtue of the supporting distance property at \bar{x} , recalling that $r_0 < r$, one obtains

$$\begin{aligned}
-\frac{\sigma}{2\kappa} & \leq \inf_{\phi \in B(\mathbf{0}, \sigma)} \inf_{x \in B(\bar{x}, r_0) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x}) - [\phi(x) - \phi(\bar{x})]}{d(x, \bar{x})} \\
& \leq \inf_{x \in B(\bar{x}, r_0) \setminus \{\bar{x}\}} \left[\frac{f(x) - f(\bar{x})}{d(x, \bar{x})} - \sup_{\phi \in B(\mathbf{0}, \sigma)} \frac{\phi(x) - \phi(\bar{x})}{d(x, \bar{x})} \right] \\
& \leq \inf_{x \in B(\bar{x}, r_0) \setminus \{\bar{x}\}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} - \inf_{x \in B(\bar{x}, r_0) \setminus \{\bar{x}\}} \sup_{\phi \in B(\mathbf{0}, \sigma)} \frac{\phi(x) - \phi(\bar{x})}{d(x, \bar{x})} \\
& \leq f^\downarrow(\bar{x}) - \frac{\sigma}{\kappa}.
\end{aligned}$$

Thus, it is $f^\downarrow(\bar{x}) \geq \sigma/2\kappa$. This completes the proof. \square

Combining Proposition 6 and Proposition 8 puts one in a position to derive the following characterization of condition (\mathcal{C}) in subdifferential terms.

Corollary 2 *Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, let $\Phi \subseteq \text{Lip}(X)/\sim$. Suppose that Φ satisfies the local supporting distance property at \bar{x} and balls in Φ are compact. Then condition (C) holds iff*

$$\mathbf{0} \in \text{int } \partial_{\Phi}^{\text{loc}} f(\bar{x}).$$

Proof The necessary part of the thesis follows from Proposition 6, after recalling that $\partial_{\Phi,0}^{\text{loc}} f(\bar{x}) \subseteq \partial_{\Phi}^{\text{loc}} f(\bar{x})$. The sufficient one comes from Proposition 8. \square

For functions enjoying a certain convexity property, it becomes possible to drop out the compactness assumption on the balls and to use the Φ -subdifferential for characterizing condition (C).

Definition 7 Let $\Phi \subseteq \text{Lip}(X)/\sim$ be a given family. A function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ is said to be Φ -convex if

$$f(x) = \sup\{\ell(x) : [\ell]_{\sim} \in \Phi, \ell \leq f\}, \quad \forall x \in X,$$

where $\ell \leq f$ means that $\ell(x) \leq f(x)$ for every $x \in X$.

Remark 6 Whenever X is a vector space, for certain families $\Phi \subseteq \text{Lip}(X)/\sim$ it happens that if the inequality $\phi(x) \leq f(x)$ holds in a neighbourhood of a point $\bar{x} \in X$, then it continues being valid on the whole space X . This is the case, for example, of Φ given by the linear functionals on X and the classic concept of convexity. Notice that the property

$$\phi(x) \leq f(x), \quad \forall x \in B(\bar{x}, r) \quad \text{implies} \quad \phi(x) \leq f(x), \quad \forall x \in X, \quad (10)$$

entails that, if f is Φ -convex, then

$$\partial_{\Phi} f(\bar{x}) = \partial_{\Phi,0}^{\text{loc}} f(\bar{x}).$$

Corollary 3 *Given a function $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, let $\Phi \subseteq \text{Lip}(X)/\sim$ satisfy the supporting distance property and property (10). Suppose that f is Φ -convex. Then condition (C) holds iff $\mathbf{0} \in \text{int } \partial_{\Phi} f(\bar{x})$.*

Proof In the light of Proposition 6 and Proposition 7, the thesis becomes an obvious consequence of Remark 6. \square

4 Consequences on stability in variational analysis

Throughout the present section, $(\mathbb{X}, \|\cdot\|)$ denotes a real Banach space, with null vector $\mathbf{0}$. Its topological dual is marked by \mathbb{X}^* , whose null vector is $\mathbf{0}^*$, whereas the duality pairing \mathbb{X}^* and \mathbb{X} is indicated by $\langle \cdot, \cdot \rangle$. Set $\mathbb{B} = B(\mathbf{0}, 1)$ and $\mathbb{S} = \{u \in \mathbb{B} : \|u\| = 1\}$ and, similarly, $\mathbb{B}^* = B(\mathbf{0}^*, 1)$ and $\mathbb{S}^* = \{u \in \mathbb{B}^* : \|u\| = 1\}$. Given a function $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \text{dom } f$, the Fréchet derivative of f at \bar{x} is denoted by $Df(\bar{x})$.

Below some situations are illustrated, in which the conditions involving the steepest descent rate discussed in Section 2 seem to be “not natural” for smooth functions.

Remark 7 (i) If function $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ is Fréchet differentiable at $\bar{x} \in \text{dom } f$, then its steepest descent rate at \bar{x} can be represented as follows

$$f^\downarrow(\bar{x}) = \inf_{u \in \mathbb{S}} \langle Df(\bar{x}), u \rangle. \quad (11)$$

Indeed, setting $o(\|x - \bar{x}\|) = f(x) - f(\bar{x}) - \langle Df(\bar{x}), x - \bar{x} \rangle$, one finds

$$\begin{aligned} \liminf_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x})}{d(x, \bar{x})} &= \liminf_{x \rightarrow \bar{x}} \left[\left\langle Df(\bar{x}), \frac{x - \bar{x}}{\|x - \bar{x}\|} \right\rangle + \frac{o(\|x - \bar{x}\|)}{\|x - \bar{x}\|} \right] \\ &= \liminf_{x \rightarrow \bar{x}} \left\langle Df(\bar{x}), \frac{x - \bar{x}}{\|x - \bar{x}\|} \right\rangle = \inf_{u \in \mathbb{S}} \langle Df(\bar{x}), u \rangle. \end{aligned}$$

So, being $Df(\bar{x}) \in \mathbb{X}^*$, in such event it must be $f^\downarrow(\bar{x}) \leq 0$. This shows that condition (C) can never be fulfilled by a function Fréchet differentiable at \bar{x} .

(ii) Condition (2) can never be satisfied by a non constant function $f \in C^1(\mathbb{X})$ admitting global minimizers. Indeed, in such case the set $\text{Argmin}(f) \neq \mathbb{X}$ is closed. Let $\bar{x} \in \text{bd Argmin}(f)$. Then there exists $(x_n)_{n \in \mathbb{N}}$, with $x_n \rightarrow \bar{x}$ as $n \rightarrow \infty$ and $x_n \notin \text{Argmin}(f)$. According to condition (2), for some $\sigma > 0$ it should be

$$f^\downarrow(x_n) = \inf_{u \in \mathbb{S}} |\langle Df(x_n), u \rangle| < -\sigma, \quad \forall n \in \mathbb{N}. \quad (12)$$

Since it is $Df(x_n) \rightarrow \mathbf{0}^* = Df(\bar{x})$ as $n \rightarrow \infty$ by continuity of the mapping $Df : \mathbb{X} \rightarrow \mathbb{X}^*$, there must exist $n_\sigma \in \mathbb{N}$ such that

$$\sup_{u \in \mathbb{S}} |\langle Df(x_n), u \rangle| = \|Df(x_n)\| \leq \frac{\sigma}{2}, \quad \forall n \in \mathbb{N}, \quad n \geq n_\sigma,$$

so one actually finds

$$\inf_{u \in \mathbb{S}} \langle Df(x_n), u \rangle \geq -\frac{\sigma}{2}, \quad \forall n \in \mathbb{N}, \quad n \geq n_\sigma$$

which contradicts evidently inequality (12).

For a nonsmooth function $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{\pm\infty\}$ the representation (11) is replaced by the side estimate

$$f^\downarrow(\bar{x}) \leq \inf_{u \in \mathbb{S}} D_{\mathcal{H}}^\downarrow f(\bar{x}; u), \quad (13)$$

where

$$D_{\mathcal{H}}^\downarrow f(\bar{x}; u) = \liminf_{\substack{v \rightarrow u \\ t \downarrow 0}} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

indicates the Hadamard lower derivative of f at \bar{x} , in the direction $u \in \mathbb{X}$ (see, for instance, [6]). While estimate (13) generally is not useful to get conditions which are sufficient for (C), it enables one to formulate the following sufficient condition for the solution existence of nondifferentiable optimization problems.

Theorem 1 *Let $f : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a l.s.c. function bounded from below. If there exists $\sigma > 0$ such that*

$$\sup_{x \in [f > \inf_X f] \cap \text{dom } f} \inf_{u \in \mathbb{S}} D_{\mathcal{H}}^{\downarrow} f(x; u) < -\sigma,$$

then $\text{Argmin}(f) \neq \emptyset$.

Proof The thesis is an obvious consequence of Proposition 3 and of the estimate (13). \square

Nevertheless, some special cases are known in which inequality (13) turns out to hold as an equality. For example, if $\mathbb{X} = \mathbb{R}^n$, by virtue of the compactness of balls, it has been shown that

$$f^{\downarrow}(\bar{x}) = \min_{u \in \mathbb{S}} D_{\mathcal{H}}^{\downarrow} f(\bar{x}; u)$$

(see [6, 26]). In this case, condition (C) can be characterized via the positivity of $D_{\mathcal{H}}^{\downarrow} f(\bar{x}; u)$ over \mathbb{S} .

On the other hand, when working with nonsmooth functions defined on Banach spaces widely tools of analysis are generalized derivative constructions based on the dual space. In this concern, observe that if $\mathfrak{A}(\mathbb{X}) \subseteq \text{Lip}(\mathbb{X})$ denotes the family consisting of all affine functions on X , then $\mathfrak{A}(\mathbb{X})/\sim$ can be identified with \mathbb{X}^* . Note that in such case $\|\cdot\|_{\text{Lip}}$ becomes the usual (uniform) norm in \mathbb{X}^* . By this choice of Φ , $\partial_{\Phi} f(\bar{x})$ coincides with the subdifferential of f at \bar{x} in the sense of convex analysis, here denoted simply by $\partial f(\bar{x})$, whereas $\partial_{\Phi}^{\text{loc}} f(\bar{x})$ coincides with the regular (aka Fréchet) subdifferential of f at \bar{x} , here denoted by $\hat{\partial} f(\bar{x})$, i.e.

$$\hat{\partial} f(\bar{x}) = \left\{ x^* \in \mathbb{X}^* : \liminf_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \right\}$$

(for detailed expositions concerning this construction see [22, 19, 23, 24]). Furthermore, whenever a l.s.c. function $f : \mathbb{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ is $\mathfrak{A}(\mathbb{X})/\sim$ -convex, then it is convex in the classical sense.

Remark 8 By standard separation arguments of convex analysis, it is not difficult to see that \mathbb{X}^* satisfies the supporting distance property. Indeed, by taking $\kappa = 1$, one actually has

$$\sup_{x^* \in \epsilon \mathbb{B}^*} \langle x^*, x - \bar{x} \rangle = \epsilon \|x - \bar{x}\|, \quad \forall x, \bar{x} \in \mathbb{X}.$$

Along with $\mathfrak{A}(\mathbb{X})$, other subclasses of $\text{Lip}(\mathbb{X})$ satisfying the supporting distance property and leading to interesting Φ -subdifferential constructions are

$$\mathfrak{S}(\mathbb{X}) = \{\phi : \mathbb{X} \longrightarrow \mathbb{R} : \phi \text{ is sublinear and continuous on } \mathbb{X}\}$$

$$\mathfrak{D}\mathfrak{S}(\mathbb{X}) = \mathfrak{S}(\mathbb{X}) - \mathfrak{S}(\mathbb{X}).$$

By specializing to a Banach space setting what established in Section 3, it is possible to extend and generalize the characterization of sharp minimality presented in [11] (Ch. 5, Lemma 3).

Theorem 2 *Let $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a l.s.c. convex function. Suppose that f is continuous at $\bar{x} \in \text{dom } f$. Then condition (C) holds iff $\mathbf{0}^* \in \text{int } \partial f(\bar{x})$.*

Proof On the base of the current subdifferential constructions, by virtue of what noticed in Remark 8, it suffices to apply Corollary 3. \square

Since a convex function is Gâteaux differentiable at a given point \bar{x} of its domain iff its subdifferential reduces to a singleton, condition (C) evidently appear to be inconsistent with such kind of smoothness at \bar{x} . Below, some of the benefits concerning the stability behaviour of nonsmooth functions are listed.

Theorem 3 *Let $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a l.s.c. convex function, which is continuous at $\bar{x} \in \text{dom } f$. Suppose that $f(\bar{x}) = 0$. If $\mathbf{0}^* \in \text{int } \partial f(\bar{x})$ the following assertions are true:*

- (i) \bar{x} is a global sharp minimizer of f ;
- (ii) \bar{x} is superstable;
- (iii) function f admits a global error bound at \bar{x} ;
- (iv) there exists $\epsilon > 0$ such that every function $\tilde{f} \in \text{Ptb}(f, \bar{x}, \epsilon)$ admits a local error bound at \bar{x} .

Proof In the light of the characterization provided by Theorem 2, assertions (i), (ii) and (iii) follow from the positivity of the steepest descent rate of f at \bar{x} (remember Proposition 1, Proposition 2, and Proposition 4). Indeed, under the current assumptions, the sharp minimality of \bar{x} can be obtained in a global form on account of inequality (9). Besides, globality of the error bound property trivially takes place because one has

$$[f \leq 0] = \{\bar{x}\} \quad \text{and} \quad f(x) = [f(x)]_+, \quad \forall x \in X.$$

As for assertion (iv), it suffices to recall Remark 4. \square

Remark 9 (i) Notice that assertion (iii) in Theorem 2 allows one to recover the implication “ $\mathbf{0}^* \in \text{int } \partial f(\bar{x}) \Rightarrow$ global error bound at \bar{x} ” appearing, among other results, in Theorem 1 of [16]. In turn, such implication enables one to derive the following well-known error bound condition for a convex inequality

$$f'(\bar{x}; v) \geq \sigma \|v\|, \quad \forall v \in \mathbb{X},$$

where $f'(\bar{x}; v)$ denotes the directional derivative of f at \bar{x} , in the direction v . Its sufficiency indeed comes directly from the dual representation

$$f'(\bar{x}; v) = \max_{x^* \in \partial f(\bar{x})} \langle x^*, v \rangle, \quad \forall v \in \mathbb{X}.$$

The necessariness instead follows from the local sharp minimality of \bar{x} , which makes the inequality

$$\frac{f(\bar{x} + tv) - f(\bar{x})}{t} \geq \sigma \|v\|$$

true for $t > 0$ small enough.

(ii) It is clear that condition $\mathbf{0}^* \in \text{int } \partial f(\bar{x})$ is far from being necessary for the error bound of f at \bar{x} . Indeed, the latter property can be achieved under the condition $\mathbf{0}^* \notin \partial f(\bar{x})$ as well (see again Theorem 1 in [16]).

In the finite-dimensional case it is possible to establish the following counterpart of Theorem 2, which applies to nonconvex functions.

Theorem 4 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$ be a function and $\bar{x} \in \text{dom } f$. Then condition (C) holds iff $\mathbf{0}^* \in \text{int } \hat{\partial} f(\bar{x})$.*

Proof Upon the identification $\mathfrak{A}(\mathbb{X})/\sim \cong \mathbb{X}^*$, all hypotheses of Corollary 2 happen to be fulfilled (recall Remark 8). Thus the thesis becomes an obvious consequence of it. \square

The consequences on the variational behaviour stability of nonconvex non-smooth functions are presented below. Notice that in the absence of convexity assertions (i) and (iii) lose their global validity.

Theorem 5 *Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$ be a function, with $\bar{x} \in \text{dom } f$. Suppose that $f(\bar{x}) = 0$. If $\mathbf{0}^* \in \text{int } \hat{\partial} f(\bar{x})$ the following assertions are true:*

- (i) \bar{x} is a local sharp minimizer of f ;
- (ii) \bar{x} is superstable;
- (iii) function f admits a local error bound at \bar{x} ;
- (iv) there exists $\epsilon > 0$ such that every function $\tilde{f} \in \text{Ptb}(f, \bar{x}, \epsilon)$ admits a local error bound.

Proof It is possible to argue as in the proof of Theorem 2, using instead the characterization provided by Theorem 4. \square

Remark 10 According to [25] a point $\bar{x} \in \mathbb{X}$ is said to give a tilt-stable local minimum of $f : \mathbb{X} \rightarrow \mathbb{R} \cup \{\pm\infty\}$, if $\bar{x} \in \text{dom } f$ and there exists $\delta > 0$ such that the set-valued mapping $\mathcal{M}_{f,\delta} : \mathbb{X}^* \rightrightarrows \mathbb{X}$ defined by

$$\mathcal{M}_{f,\delta}(x^*) = \left\{ y \in \mathbb{X} : y \text{ solves } \min_{x \in B(\bar{x}, \delta)} [f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle] \right\}$$

is single-valued and Lipschitz on some neighbourhood of $\mathbf{0}^*$, with $\mathcal{M}_{f,\delta}(\mathbf{0}^*) = \{\bar{x}\}$. Since for any $x^* \in \mathbb{X}^*$ and $x \in \mathbb{X}$, one has $|\nabla x^*|(x) = \|x^*\|$, from the superstability of \bar{x} for f follows that, if taking $\delta < f^\downarrow(\bar{x})$, the mapping $\mathcal{M}_{f,\delta}$ in a neighbourhood of $\mathbf{0}^*$ constantly takes the value $\{\bar{x}\}$. Thus, Theorem 3 and Theorem 5 entail also the tilt-stability of \bar{x} .

5 Conclusions

The meditations exposed in the present paper should convince a reader that in optimization nonsmoothness does not mean necessarily a pathology, leading only to handicaps in the problem analysis. In certain situations having to do with the solution stability and the sufficiency of optimality conditions, nonsmoothness affords a robustness behaviour that smoothness can not guarantee. Here some evidences of such a phenomenon are collected and discussed. An aspect which seems to be remarkable is that a unifying study of the issue can be conducted already in a metric space setting, via a positivity condition on the steepest descent rate.

References

1. *Correspondance d'Hermite et de Stieltjes: Tome I-II*, Gauthier-Villars, Paris (1904-1905)
2. MARINO, A., *Punti stazionari e curve di massima pendenza in domini non convessi. Geodetiche con ostacolo*, Proceedings of the Congress "Studio di problemi-limite della analisi funzionale", Bressanone 1981, Pitagora Editrice, Bologna, 129–154 (1982) [in Italian]
3. GIANNESI, F., *Semidifferentiable functions and necessary optimality conditions*, J. Optim. Theory Appl. **60**, no. 2, 191–241 (1989)
4. DEMYANOV, V.F., *Conditions for an extremum in metric spaces*, J. Global Optim. **17**, no. 1-4, 55–63 (2000)
5. DEMYANOV, V.F., *An old problem and new tools*, Optim. Meth. Soft. **20**, no. 1, 53–70 (2005)
6. DEMYANOV, V.F., *Conditions for an extremum and calculus of variations*, Vyshaya Shkola, Moscow (2005) [in Russian]
7. DEMYANOV, V.F., *Nonsmooth optimization*, in Nonlinear optimization, Lecture Notes in Math. 1989, Springer, Berlin, 55–163 (2010)
8. ZASLAVSKI, A.J., *An exact penalty approach to constrained minimization problems on metric spaces*, Optim. Lett. **7**, no. 5, 1009–1016 (2013)
9. ZASLAVSKI, A.J., *An approximate exact penalty in constrained vector optimization on metric spaces* J. Optim. Theory Appl. **162**, no. 2, 649–664 (2014)
10. DE GIORGI, E., MARINO, A., and TOSQUES, M., *Problems of evolution in metric spaces and maximal decreasing curves*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **68**, no. 3, 180–187 (1980) [in Italian]
11. POLYAK, B.T., *Introduction to optimization*, Optimization Software, New York (1987)
12. ROCKAFELLAR, R.T., *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14**, 877–898 (1976)
13. FERRIS, M.C., *Finite termination of the proximal point algorithm*, Math. Programming **50**, no. 3, (Ser. A), 359–366 (1991)
14. BURKE, J.V. and FERRIS, M.C., *Weak sharp minima in mathematical programming*, SIAM J. Control Optim. **31**, no. 5, 1340–1359 (1993)
15. BURKE, J.V. and DENG, S., *Weak sharp minima revisited, part I: basic theory, well-posedness in optimization and related topics*, Control Cybernet. **31**, no. 3, 439–469 (2002)
16. KRUGER, A., NGAI, H.V., and THÉRA, M., *Stability of error bounds for convex constraint systems in Banach spaces*, SIAM J. Optim. **20**, no. 6, 3280–3296 (2010)
17. FACCHINEL, F. and PANG, J.-S., *Finite-dimensional variational inequalities and complementarity problems. Vol. I*, Springer Series in Operations Research. Springer-Verlag, New York (2003)
18. FABIAN, M.J., HENRION, R., KRUGER, A.Y., and OUTRATA, J.V., *Error bounds: necessary and sufficient conditions*, Set-Valued Var. Anal. **18**, no. 2, 121–149 (2010)
19. PENOT, J.-P., *Calculus without derivatives*, GTM, 266. Springer, New York (2013)

-
20. KRUGER, A.Y., *Error bounds and metric regularity*, to appear in Optimization, DOI: 10.1080/02331934.2014.938074
 21. PALLASCHKE, D. and ROLEWICZ, S., *Foundations of Mathematical Optimization*, Kluwer Academic Publishers, Dordrecht (1997)
 22. MORDUKHOVICH, B.S., *Variational analysis and generalized differentiation. I. Basic theory*. Fundamental Principles of Mathematical Sciences, 330. Springer-Verlag, Berlin (2006)
 23. ROCKAFELLAR R.T. and WETS, R. J.-B., *Variational analysis*. Fundamental Principles of Mathematical Sciences, 317. Springer-Verlag, Berlin (1998)
 24. SCHIROTZEK, W., *Nonsmooth analysis*, Universitext. Springer, Berlin (2007)
 25. POLIQUIN, R.A. and ROCKAFELLAR R.T., *Tilt stability of a local minimum*, SIAM J. Optim. **8**, no. 2, 287–299 (1998)
 26. PAPPALARDO, M. and UDERZO, A., *G-semidifferentiability in Euclidean spaces*, J. Optim. Theory Appl. **101**, no. 1, 221–229 (1999)